

كيف يمكن لتطبيقات الذكاء الاصطناعي أن تكون

مُتَحَيِّزة؟

كيف يعمل الذكاء الاصطناعي؟

كيف يمكن للذكاء الاصطناعي أن يكون مُتَحَيِّزاً؟

أزمة اختيار موظفي أمازون

تقييم المخاطر الجنائية

على الرغم من التطور الضخم الذي يساهم فيه الذكاء الاصطناعي في الكثير من المجالات، إلا أن واحدة من أهم الاشكاليات التي ظهرت في السنوات السابقة هي عنصرية بعض تطبيقات الذكاء الاصطناعي على أساس جنس أو عرق معين؛ للدرجة التي تجعل بعض التطبيقات تنتهك حقوق الأفراد وتتحكم في مسار حياتهم. كانت مسار – مجتمع التقنية والقانون قد نشرت حلقتين على بودكاست كُونَكْت تتناول تحيُّز وعنصرية بعض تطبيقات وخوارزميات الذكاء الإصطناعي.

كيف يعمل الذكاء الاصطناعي؟

تُوفّر تقنيات الذكاء الاصطناعي محاكاة الذكاء البشري للحواسيب، وله أساليب ونظريات رئيسية مثل؛ التفكير البشري والتفكير العقلاني والتصرف البشري والتصرف العقلاني. وبشكل عام، يعمل الذكاء الاصطناعي بواسطة جمع كميات كبيرة من البيانات و تغذيته بها، من ثم معالجتها عبر خوارزمياته الذكية، مما يسمح للبرنامج أو الآلة بالتعلم من الأنماط والخصائص الموجودة بالبيانات أو المعلومات المُستخرجة منها. هذا التعليم يُستخدم لاحقًا من قبل الآلة لتطوير قدراتها في التعامل مع المهام المُعدّة لها، على سبيل المثال، يمكن لتطبيق دردشة قائم على الذكاء الاصطناعي أن يُغذّي بأمثلة من الدردشات النصية؛ من ثمّ يتعلم إنتاج ردود قريبة من البشرية مع الأشخاص الذين يتحدثون إليه. مثال آخر، يمكن لأداة التعرف على الصور أن تتعلم التعرف على الأشخاص أو الأشياء الموجودة بالصور ووصفها.

يتضمن الذكاء الاصطناعي العديد من **النظريات والمنهجيات**، ويُمكن الإشارة إلى **المجالات الأكثر انتشارًا**، كالتالي:

- **التعلّم الآلي**: أو التعلم بالتجربة، حيث يوفر لأنظمة الحواسيب القدرة على التعلم والتطوير تلقائيًا بواسطة التجربة، ويكون ذلك عبر تطوير خوارزميات يُمكنها تحليل البيانات وخلق تكهّنات بناءً على معالجة البيانات. ويُستخدم التعلّم الآلي على نطاق واسع في العديد من المجالات مثل الرعاية الصحية وصناعة الأدوية، حتى العديد من تطبيقات الترفيه كترشيحات أفلام نتفليكس الملائمة للمستخدم.
- **التعلم العميق**: أو التعلّم الذاتي، حيث تعتمد تطبيقات الذكاء الاصطناعي على الشبكات العصبية الاصطناعية، وتطوّر الخوارزميات بحيث يكون للآلة القدرة على تعليم نفسها عبر محاكاة الخلايا العصبية البشرية في تحليل البيانات واتخاذ القرارات، وتكون نتائج تصرفات الآلة قريبة من الفعل البشري. تطبيقات مثل Siri و Alexa من التطبيقات التي تستخدم التعلم العميق.
- **الشبكات العصبية**: كما ذكرنا من قبل تُستخدم في التعلم العميق. وهي أنظمة للحواسيب على غرار الشبكة العصبية في الدماغ البشري. وتتعلم تلك الأنظمة من تحليل كم ضخم من البيانات عدة مرات لإيجاد وتحديد أنماط أو ارتباطات غير

مُحدّدة مسبقًا، فتكون الآلة قادرة على إخراج نتيجة واحدة من عدّة مُدخلات. على سبيل المثال، تدريب الآلة على وجود كائن مُعين في الصور عبر إدخال كم ضخم من الصور التي تُظهر هذا الكائن ومن ثم تكون الآلة قادرة لاحقًا على تمييز وجود هذا الكائن في الصور.

- الحوسبة المعرفية: تسعى إلى إعادة إنشاء ومحاكاة عملية التفكير البشري في نموذج حاسوبي، كفهم لغة بشرية أو معنى للصور مثلًا. تعمل الحوسبة المعرفية والذكاء الاصطناعي معًا على منح الآلات سلوكيات شبيهة بالإنسان وقدراته على معالجة المعلومات، فيُحسّن ويُطوّر التفاعل بين البشر والآلة.
- معالجة اللغات الطبيعية: حيث تكون الآلة قادرة على معالجة اللغات البشرية الطبيعية وتفسيرها واستخدامها، فتكون الآلة قادرة على التفاعل مع البشر مع فهم للسياق وإنتاج استجابات منطقية. مثال على ذلك، مترجم سكايب، الذي يفسر الكلام بلغات متعددة في الوقت الفعلي لتسهيل التواصل.
- الرؤية الحاسوبية: وهي تقنية تستخدم التعلم العميق لتحديد ومعالجة وتفسير البيانات البصرية؛ بما في ذلك الصور والرسوم البيانية والجدول ووثائق PDF، فضلًا عن النصوص والفيديو.

كيف يمكن للذكاء الاصطناعي أن يكون مُتحيّزًا؟

بشكل أساسي تكون عنصرية الذكاء الاصطناعي بسبب البيانات المُتحيّزة التي يتم تدريب الذكاء الاصطناعي بواسطتها، أيضًا وجود افتراضات عنصرية وانعكاس لمشكلات المجتمع العنصرية في خوارزميات الذكاء الاصطناعي؛ تجعله يتصرف بطرق تعكس عدم التسامح أو التمييز الموجود في المجتمع.

أزمة اختيار موظفي أمازون

أحد الأمثلة الشهيرة على تحيز الذكاء الاصطناعي كانت شركة "أمازون" التي قامت في عام 2014 ببناء نظام ذكاء اصطناعي للمساعدة في عملية التوظيف بالشركة، لتجد أن النظام الجديد يميز ضد النساء.

شكّلت "أمازون" فريقاً من المبرمجين للعمل على جعل عملية اختيار الموظفين أوتوماتيكية، هذا الفريق غذى نظام الذكاء الاصطناعي بمجموعة كبيرة من السير الذاتية التي قُدمت إلى أمازون على مدار 10 سنوات، وكان من المفترض أن يقوم النظام -بناءً على البيانات التي تم تغذيته بها- بتقديم توصيات واختيار أفضل المرشحين للوظائف الجديدة.

بعد مرور عام، لاحظ المهندسون أن النظام يميز ضد النساء، ومع مراجعة السير الذاتية التي تم تغذية النظام بها لتدريبه على اختيار الأنسب للوظائف، اكتشفت الشركة أن أغلب هذه السير الذاتية كانت لذكور، وبالتالي أعطى نظام الذكاء الاصطناعي تقييماً منخفضاً للسير الذاتية التي تحتوي على أسماء مؤنث، أو بيانات أخرى مثل ممارسة رياضة معينة.

تقييم المخاطر الجنائية

مثال آخر على تحيز الذكاء الاصطناعي حدث في العام ذاته، 2014؛ عندما وجدت فتاة أمريكية تُدعى "بريشا بوردون" وصديقة لها دراجة و"سكوتر" غير مقفلين في الشارع وقامت باستخدامهما. لكن الفتاتان البالغان من العمر 18 عاماً أدركتا أن الدراجة والسكوتر صغيران جداً فتركاهما في الشارع وانصرفا، في الوقت الذي جاءت فيه امرأة تركض وراءهن قائلة "هذه أغراض طفلي". أحد الجيران الذي شهد الواقعة كان قد أبلغ الشرطة بالفعل، وتم القبض على بريشا وصديقتها ووجهت إليهما تهمة السطو وسرقة أشياء تقدر قيمتها الإجمالية بـ 80 دولاراً. "بريشا" كان لديها سجلا إجراميا يتعلق بجنح كانت قد ارتكبتها قبل أن تبلغ السن القانوني.

في العام السابق لتلك الواقعة، ألقت الشرطة القبض على شخص يبلغ من العمر 41 عاماً يدعى "فيرنون براتر" بسبب سرقة أدوات بقيمة 86.35 دولاراً من أحد المتاجر. لم تكن تلك هي الجريمة الأولى لـ "براتر"، حيث إنه كان قد أدين في السابق بالسطو المسلح، وقضى خمس سنوات في السجن.

خلال المحاكمة، تم استخدام نظام ذكاء اصطناعي لعمل ما يعرف بـ "تقييم المخاطر" لكل منهما، والتنبؤ باحتمالية ارتكابهما للمزيد من الجرائم في المستقبل، وكذلك لإصدار قرارات حول إمكانية إطلاق سراح المتهمين في مرحلة من مراحل نظام العدالة الجنائية. "بريشا"،

والتي كانت سمراء البشرة، صنفتها النظام على أنها ذات خطورة كبيرة. أما "فيرنون"، وهو أبيض البشرة، تم تصنيفه بأنه منخفض الخطورة. بعد عامين من المحاكمة، لم يتم اتهام "بريشا" بارتكاب أي جرائم جديدة، فيما كان "فيرنون" يقضي عقوبة بالسجن لمدة ثماني سنوات لاقتحام مستودع وسرقة أجهزة إلكترونية بقيمة آلاف الدولارات.

تحيز نظام الذكاء الاصطناعي ضد "بريشا" وغيرها من ذوي البشرة السمراء جاء نتيجة تحيز البيانات التي تم تغذيته بها ضدهم، والتي تفيد بأنهم لديهم احتمالية أكبر لارتكاب الجرائم، ولذلك فإن معدل تصنيفهم على أنهم أكثر خطورة ضعف معدل تصنيف المتهمين ذوي البشرة البيضاء الذي يتم تصنيفهم على أنهم منخفضو الخطورة. أي أن تحيز الذكاء الاصطناعي يحدث عندما تكون البيانات نفسها متحيزة.

تكمن مخاطر تحيز الذكاء الاصطناعي في تفاقم التمييز القائم على النوع الاجتماعي وتعزيز القوالب النمطية الجنسانية والعرقية، وغيرها من المخاطر، وهو ما يمكن أن يتجسد في سوق العمل والتعليم وإعلانات الإنترنت ووسائل التواصل الاجتماعي والضرائب ونظام العدالة. لذلك، هناك حاجة إلى اختبار وتقييم أنظمة الذكاء الاصطناعي وتغذيتها ببيانات متوازنة لتقليل التحيز وتحسين دقة الأنظمة.